Advances in Applied Mathematics and Mechanics
Adv. Appl. Math. Mech., Vol. **6**, No. 2, pp. 145-164

DOI: 10.4208/aamm.2013.m175
April 2014

An Efficient Proximity Point Algorithm for Total-Variation-Based Image Restoration

Wei Zhu[1], Shi Shu[2,*] and Lizhi Cheng[3]

[1] *School of Mathematics and Computational Science, Xiangtan University, Xiangtan 411105, Hunan, China*
[2] *Hunan Key Laboratory for Computation and Simulation in Science and Engineering, Xiangtan University, Xiangtan 411105, Hunan, China*
[3] *Department of Mathematics and Computational Science, College of Science, National University of Defense Technology, Changsha 410073, Hunan, China*

Received 30 March 2013; Accepted (in revised version) 8 August 2013

Available online 2 April 2014

Abstract. In this paper, we propose a fast proximity point algorithm and apply it to total variation (TV) based image restoration. The novel method is derived from the idea of establishing a general proximity point operator framework based on which new first-order schemes for total variation (TV) based image restoration have been proposed. Many current algorithms for TV-based image restoration, such as Chambolle's projection algorithm, the split Bregman algorithm, the Bermúdez-Moreno algorithm, the Jia-Zhao denoising algorithm, and the fixed point algorithm, can be viewed as special cases of the new first-order schemes. Moreover, the convergence of the new algorithm has been analyzed at length. Finally, we make comparisons with the split Bregman algorithm which is one of the best algorithms for solving TV-based image restoration at present. Numerical experiments illustrate the efficiency of the proposed algorithms.

AMS subject classifications: 68U10, 65F22, 65K10

Key words: Proximity point operator, image restoration, total variation, first-order schemes.

1 Introduction

The well known Rudin-Osher-Fatemi (ROF) total variation (TV) model [1] was introduced to image restoration in 1992 by Rudin and Osher et al., and gained a great number of studied interest and applications [2, 3] such as image deblurring, image inpainting

*Corresponding author.
Email: zw_1012@163.com (W. Zhu), shushi@xtu.edu.cn (S. Shu), clzcheng@vip.sina.com (L. Cheng)

during the last over twenty years. The central difficulty in the TV model lies in the non-differentiability of the total variation norm and the large dimension of the underlying image from a numerical point of view.

Up to now, there are a wide variety of methods to address the total variation norm minimization in image proceeding. We here just list a few of them: Chambolle's projection algorithm [4, 5], the split Bregman algorithm [6, 7], Nestrov's schemes [8, 9], the Bermúdez-Moreno algorithm [10] and the FP^2O algorithm [11]. We would like to explain why we are interested in these methods: Chambolle's projection algorithm has grown very popular since it is the first algorithm minimizing the ROF model exactly; both the split Bregman method and Nestrov's schemes gain lots of studying interest in many cases such as frame-based image restoration [12], segmentation and surface reconstruction [13], and sparse recovery [14–16]; the Bermúdez-Moreno and the FP^2O algorithms are most recently proposed, and the Bermúdez-Moreno's algorithm is comparable with Nestrov's schemes while the FP^2O algorithm can be viewed as a modification of the split Bregman algorithm.

When the methods above are applied to the total variation denoising, many of them have a common numerical scheme via some modifications. For example, Aujol in [17] showed that the modification of Chambolle's projection algorithm has the exact scheme of the Bermúdez-Moreno algorithm, Micchell et al. in [11] illustrated that with some modification the split Bregman algorithm reduces to the Jia-Zhao denoising algorithm [18] which can be considered as a special case of the FP^2O algorithm. The first main contribution of the paper is of discovering the connection between the Bermúdez-Moreno algorithm and the FP^2O algorithm; both of them can be viewed as generation of the Chambolle's projection algorithm but from different angles: the FP^2O is based on the Picard sequence, and the Bermúdez-Moreno algorithm is to extend the operator in square term from identical operator to symmetric positive operator. Under the proximity point operator frameworks, we firstly extend the FP^2O algorithm from image denoising to image deconvolution and image restoration based on wavelet with total variation, and introduce new and efficient schemes. That is the second main contribution of the paper. With these derived schemes we conclude that the algorithms above are all special cases of our proposed schemes. Moreover, we also prove the convergence of the proposed schemes by introducing the Opial κ-averaged property [19]. Since Aujol in [17] have tested numerous numerical comparisons between the Bermúdez-Moreno algorithm and the Nestrov's schemes, we have decided to make some comparisons with the split Bregman method to test the efficiency of the proposed schemes.

Before presenting the plan of the paper, we emphasis once the main contributions of the paper:

- Discovering the connection between the Bermúdez-Moreno algorithm and the FP^2O algorithm; both of them can be viewed as generation of the Chambolle's projection algorithm but from different angles.
- Extending the FP^2O algorithm from image denoising to image deconvolution and

image restoration based on wavelet with total variation, and introducing new and efficient schemes for them. Many current algorithms can be viewed as special cases of such schemes.

- Presenting numerical comparisons with the split Bregman algorithm.

The remainder of the paper is organized as follows. In Section 2, we formulate the TV-based image deconvolution model (TVL2) and the image restoration model based on wavelet analysis with TV (TVL1L2) by restricting our attention onto the discrete setting. In Section 3, we introduce the proximity point operator theory, and then propose fixed point formulations for the TVL2 and the TVL1L2 models. In Section 4, based on the fixed point formulations, we design new iteration schemes, and provide fast algorithms by using the fast Fourier transformation. The relationship with other algorithms is also discussed. In Section 5, we strictly prove the convergence of the new algorithms by introducing the Opial κ-averaged property [19]. In Section 6, we illustrate our study by numerous comparisons with the split Bregman algorithm.

2 Problem formulation

Image restoration is one of the earliest and most classical linear inverse problems in image processing. In this class of problems, one wants to recover an original image u from its degraded observation f. In other words, we want to recover u by solving the following linear inverse problem

$$f=Ku+\eta, \tag{2.1}$$

where K represents a blurring or convolution operator for image deblurring and an identity operator for image denoising, and η is additive noise. For simplicity, we assume that the underlying image has square domains and adopt the vector notation for image, where the pixels on an $n\times n$ image are stacked into an n^2-vector, e.g., lexicographic order. Thus our discussed space can be fixed as $\mathbb{R}^{n^2}$. Given f and K, the TVL2 model for image restoration is

$$\min \frac{1}{2}\|Ku-f\|_2^2+\beta TV(u), \tag{2.2}$$

with β being a regularization parameter that reflects the noisy level. The discrete total variation norm of u has two forms: one is the isotropic TV norm given by

$$TV_1(u)=\sum_{i=1}^{n^2}\|D_iu\|_2, \tag{2.3}$$

where $D_iu\in\mathbb{R}^2$ denotes the discrete gradient of f at pixel i; the other is the anisotropic TV norm given by

$$TV_2(u)=\sum_{i=1}^{n^2}\|D_iu\|_1. \tag{2.4}$$

Let $D^{(j)}\in\mathbb{R}^{n^2\times n^2}$, $(j=1,2)$ be the first order finite difference matrices in the horizontal and vertical directions, respectively. Then $D_i\in\mathbb{R}^{n^2\times n^2}$ is a two-row matrix formed by stacking the i-th rows of $D^{(1)}$ and $D^{(2)}$. We stack $D^{(1)}$ and $D^{(2)}$ into a matrix as

$$B=\left(D^{(1)^T},D^{(2)^T}\right)^T\triangleq(D^{(1)};D^{(2)}).$$

In order to describe the TV norm as a composition of a convex function with a linear transformation, we define two functions as follows. $\psi_1:\mathbb{R}^{2n^2}\to\mathbb{R}$ as

$$\psi_1(x):=\sum_{i=1}^{n^2}\|(x_i;x_{n^2+i})\|_2, \tag{2.5}$$

and $\psi_2:\mathbb{R}^{2n^2}\to\mathbb{R}$ as

$$\psi_2(x):=\sum_{i=1}^{n^2}\|x\|_1. \tag{2.6}$$

Then we have $TV_1(u)=\psi_1(Bu)$ and $TV_2(u)=\psi_2(Bu)$. With these preparations, we write the TVL2 model for image restoration in a general form:

$$\min\frac{1}{2}\langle Au,u\rangle-\langle g,u\rangle+\varphi(Bu), \tag{2.7}$$

with $A=K^TK+\alpha I$, $g=K^Tf$ and $\varphi(\cdot)$ a convex function. This framework has been used by J. F. Aujol in [17] for studying the Bermúdez-Moreno algorithm. We start our discussion with such a framework since many image processing problems can be formulated into it, and at the same time, it is convenient to build the connection between the Bermúdez-Moreno algorithm and the FP^{2}O algorithm. Here and from now on, we assumed that the matrix A is symmetric positive definite and hence the objective function in (2.7) is strongly convex. Though in practice the matrix A is not in such a case, we can make up it by adding a square term $\alpha\|u\|^2$ in the TVL2 model which has little effect on the objective function as the parameter α tends to zero. If we choose the matrix $A=I$, then the TVL2 model becomes the TV denoising model and the problem (2.7) is equivalent to computing the proximity point of the degraded image f; this has been studied by C. A. Micchelli in [11]. The difference here is that we assume the matrix to be symmetric positive definite which will cover more applications in image processing. Except this, we extend the framework to a more complicated case; that is

$$\min\frac{1}{2}\langle Au,u\rangle-\langle g,u\rangle+\varphi_1(Bu)+\varphi_2(Wu), \tag{2.8}$$

where $\varphi_i(\cdot)$, $i=1,2$, are two convex functions and W is some wavelet frame [20]. Many image processing models can be written as the form of (2.8). For example, the TVL1T2 model whose objective function is

$$\frac{1}{2}\|Ku-f\|_2^2+\mu_1TV(u)+\mu_2\|Wu\|_1$$

is a case of choosing $\varphi_i(\cdot)$, $i=1,2$ to be the total variation norm and the 1-norm respectively. This model also arises in the magnetic resonance (MR) image reconstruction problem [21] in which wavelet analysis with TV performs better than the case of that with a single regularized term. Up to now, there are a few algorithms for dealing with this model [6, 22, 23]. S. Becker in [24] even pointed out that there might be no first-order algorithms that can deal with complicated objectives like $f(u)=TV(u)+\|Wu\|_1$ for a non-diagonal and non-orthogonal W before their paper. In the paper, we will show that we can design first-order algorithms for such a complicated composite objective under the proximity point framework. Additionally, we will make comparisons with the split Bregman algorithm to illustrate our proposed algorithm performs better.

3 Proposed method under the proximity operator framework

3.1 Proximity operator

The foundational theory of proximity operators was introduced by Moreau [25] in 1962. Recently, P. Combettes et al. [26–28] developed a series of theory on proximity operators and applied them to signal recovery problems. Let $\mathcal{H}$ be a real Hilbert space with inner product $\langle\cdot|\cdot\rangle$ and a corresponding norm $\|\cdot\|$, and $\Gamma_0(\mathcal{H})$ be a class of lower semicontinuous, convex, and proper functions from $\mathcal{H}$ to $(-\infty,+\infty]$. The proximity operator of a function $f\in\Gamma_0(\mathcal{H})$ is the operator $prox_f:\mathcal{H}\to\mathcal{H}$ which maps every $x\in\mathcal{H}$ to the unique minimizer of the function $f+\|x-\cdot\|^2/2$, i.e., for $\forall x\in\mathcal{H}$,

$$prox_f x := \arg\min_{y\in\mathcal{H}} f(y)+\frac{1}{2}\|x-y\|^2. \tag{3.1}$$

We will calculate the proximity operators ψ_1 and ψ_2. From their expression, it is sufficient to present two examples; they are one-dimensional function $f_1=\|\cdot\|_1$ and two-dimensional function $f_2(x,y)=\sqrt{x^2+y^2}$ whose proximity is given by

$$prox_{\frac{1}{\lambda}f_1} x := \max\Big(\|x\|_1-\frac{1}{\lambda},0\Big)\,\mathrm{sign}(x), \tag{3.2a}$$

$$prox_{\frac{1}{\lambda}f_2} u := \max\Big(\|u\|_2-\frac{1}{\lambda},0\Big)\frac{u}{\|u\|_2}. \tag{3.2b}$$

Another important concept is the subdifferential of a convex function defined as follows:

Definition 3.1 (see [29]). Let $f:\mathcal{H}\to\mathbb{R}$ be a convex function. A variable v is a subgradient of f at $u\in\mathcal{H}$ if

$$f(x)\geq f(u)+\langle v,x-u\rangle, \quad \forall x\in\mathcal{H}.$$

The set of all subgradients of f at $u\in\mathcal{H}$ is called the subdifferential of f at $u\in\mathcal{H}$, denoted by $\partial_f(u)$.

In the paper, we need the following result that characterizes the relationship between the proximity operator and the subdifferential of a convex function.

Lemma 3.1 (see [12])**.** *If φ is a convex function on $\mathcal{H}$ and $u\in\mathcal{H}$, then $v\in\partial_{\varphi}(u)$ if and only if $u=prox_{\varphi}(u+v)$.*

From the lemma above, every element v, namely a subgradient in the subdifferential φ at point u, can be formulated into an equation by the proximity point operator at $u+v$. Such a property will play a central role in our algorithm designing. Also because of the relationship between the proximity operator and the subdifferential of a convex function, we call our work the proximity point operator framework.

3.2 Fixed point formulation for TVL2 model

Since the objective function in (2.7) has been assumed to be strictly convex, the unique solution can be characterized by

$$0\in Au-g+\lambda B^T\circ\partial\frac{1}{\lambda}\varphi\circ B(u), \tag{3.3}$$

which is a necessary and sufficient condition for a variable to be a solution of (2.7), where $\lambda>0$ is a parameter for convergence analysis. We have observed that every element in the subdifferential φ/λ at point Bu can be characterized. If we introduced a variable v belong to the subdifferential φ/λ at point Bu to satisfy the condition (3.3), and then we can describe that condition as two equations, given by

$$0=Au-g+\lambda B^Tv, \tag{3.4a}$$
$$Bu=prox_{\frac{1}{\lambda}\varphi}(Bu+v). \tag{3.4b}$$

We have assumed that the matrix A is symmetric positive definite and hence invertible, so $u=A^{-1}(g-\lambda B^Tv)$. Replacing the expression of u into the other equation above, we get

$$BA^{-1}(g-\lambda B^Tv)=prox_{\frac{1}{\lambda}\varphi}\big(BA^{-1}(g-\lambda B^Tv)+v\big). \tag{3.5}$$

It follows that

$$v+BA^{-1}(g-\lambda B^Tv)=v+prox_{\frac{1}{\lambda}\varphi}\big(BA^{-1}(g-\lambda B^Tv)+v\big). \tag{3.6}$$

We introduce the affine transformation $L:\mathbb{R}^{2n^2}\to\mathbb{R}^{2n^2}$ given by

$$Lv:=BA^{-1}g+(I-\lambda BA^{-1}B^T)v, \tag{3.7}$$

and the operator $S:\mathbb{R}^{2n^2}\to\mathbb{R}^{n^2}$,

$$S:=\big(I-prox_{\frac{1}{\lambda}\varphi}\big)\circ L. \tag{3.8}$$

Then we can write the Eq. (3.6) as $v=Sv$. With these discussions, we conclude that the necessary and sufficient condition for a variable to be a solution of (2.7) can reduce to

$$v=Sv, \tag{3.9a}$$
$$u=A^{-1}(g-\lambda B^T v). \tag{3.9b}$$

If we can obtain a fixed point of the operator S, then the minimizer of (2.7) can be got at once. In the next section, we will design an iteration algorithm to generate a minimizer based on the above equations.

3.3 Fixed point formulation for TVL1L2 model

Following the thought of the deriving the fixed point formulation for TVL2 model, we derive a similar formulation for the TVL1L2 model in this subsection. At first, we define some operators as follows. Affine transformations $L_1:\mathbb{R}^{3n^2}\to\mathbb{R}^{2n^2}$ defined as

$$L_1\circ(v_1;v_2):=BA^{-1}g+[I-\lambda BA^{-1}B^T-\lambda BA^{-1}W^T](v_1;v_2), \tag{3.10}$$

and $L_2:\mathbb{R}^{3n^2}\to\mathbb{R}^{n^2}$ as

$$L_2\circ(v_1;v_2):=WA^{-1}g+[I-\lambda WA^{-1}B^T-\lambda WA^{-1}W^T](v_1;v_2), \tag{3.11}$$

where $v_1\in\mathbb{R}^{2n^2}$, $v_2\in\mathbb{R}^{n^2}$. We define composition operators $S_1:\mathbb{R}^{3n^2}\to\mathbb{R}^{2n^2}$,

$$S_1:=\big(I-prox_{\frac{1}{\lambda}\varphi_1}\big)\circ L_1, \tag{3.12}$$

and $S_2:\mathbb{R}^{3n^2}\to\mathbb{R}^{n^2}$,

$$S_2:=\big(I-prox_{\frac{1}{\lambda}\varphi_2}\big)\circ L_2, \tag{3.13}$$

and $\bar{S}:\mathbb{R}^{3n^2}\to\mathbb{R}^{3n^2}$,

$$\bar{S}\circ(v_1;v_2):=(S_1\circ(v_1;v_2);S_2\circ(v_1;v_2)), \tag{3.14}$$

where we let $a=(a_1;a_2)\triangleq(a_1,a_2)^T$, $v=(v_1;v_2)\triangleq(v_1^T,v_2^T)^T$, $A=(A_1;A_2)\triangleq(A_1^T,A_2^T)^T$ for scalars a_i, vectors v_i and matrices A_i for appropriate sizes. We formulate our conclusion into the following proposition for the TVL1L2 model.

Proposition 3.1. Let λ be a positive number, then $u=A^{-1}(g-\lambda B^T v_1-\lambda W^T v_2)$ is the unique solution of (2.8) if and only if $(v_1;v_2)$ is a fixed point of $\bar{S}$.

Proof. Note that the objective function of (2.8) is strictly convex, we derive a necessary and sufficient condition for a variable to be the solution of (2.8), that is

$$0\in Au-g+B^T\circ\partial\varphi_1\circ B(u)+W^T\circ\partial\varphi_2\circ W(u). \tag{3.15}$$

Thus there exist subgradients $v_1 \in \partial_{\varphi_1/\lambda} \circ B(u)$ and $v_2 \in \partial_{\varphi_2/\lambda} \circ W(u)$ such that the solution $u = A^{-1}(g - \lambda B^T v_1 - \lambda W^T v_2)$ to be the unique solution of (2.8). By Lemma 3.1, we have that

$$Bu = prox_{\frac{1}{\lambda}\varphi_1}(v_1 + Bu), \tag{3.16a}$$

$$Wu = prox_{\frac{1}{\lambda}\varphi_2}(v_2 + Wu). \tag{3.16b}$$

Utilizing the operators defined above together with the expression of u, we get

$$v_1 + Bu = L_1 \circ (v_1; v_2), \tag{3.17a}$$

$$v_2 + Wu = L_2 \circ (v_1; v_2). \tag{3.17b}$$

Therefore, it follows that

$$L_i \circ (v_1; v_2) = v_i + prox_{\frac{1}{\lambda}\varphi_i} \circ L_i \circ (v_1; v_2), \quad i = 1,2. \tag{3.18}$$

i.e., $S_i \circ (v_1; v_2) = v_i$, $i = 1,2$. We finally get $\bar{S} \circ (v_1; v_2) = (v_1; v_2)$. Now, we can write the necessary and sufficient condition (3.15) as the following equations:

$$u = A^{-1}(g - \lambda B^T v_1 - \lambda W^T v_2), \tag{3.19a}$$

$$(v_1; v_2) = \bar{S} \circ (v_1; v_2), \tag{3.19b}$$

which concludes the proposition. □

4 Iterative algorithms based on fixed point formulations

4.1 Iterative algorithms for the TVL2 and TVL1L2 models

In this subsection, we will firstly introduce concepts such as the Picard sequence and κ-averaged operator. Then with these tools we will construct iterative algorithms to compute fixed points of the operator S and $\bar{S}$.

For a given $v^0 \in \mathbb{R}^{n^2}$ and an operator $P : \mathbb{R}^{n^2} \to \mathbb{R}^{n^2}$, we define $v^{n+1} := P(v^n)$. The sequence $\{v^n\}$ is called the Picard sequence of the operator P. For any $\kappa \in (0,1)$, the κ-averaged operator P_κ of P is defined by

$$P_\kappa := \kappa I + (1-\kappa)P. \tag{4.1}$$

Since

$$Sv = \big(I - prox_{\frac{1}{\lambda}\varphi}\big)\big(BA^{-1}(g - \lambda B^T v) + v\big), \tag{4.2}$$

if we let $u^m = A^{-1}(g - \lambda B^T v^m)$, then the Picard sequence of the κ-averaged operator S_κ is given by

$$v^{m+1} = \kappa v^m + (1-\kappa)\big(I - prox_{\frac{1}{\lambda}\varphi}\big)(Bu^m + v^m). \tag{4.3}$$

In the following Section we will prove that the Picard sequence $\{v^m\}$ converges to a fixed point of S and hence conclude that $\{u^m\}$ converges to the solution of (2.7) or the TVL2 model. We list the completion iterative algorithm as follows:

$$\begin{cases} u^m = A^{-1}(g-\lambda B^T v^m), \\ v^{m+1} = \kappa v^m + (1-\kappa)\big(I - prox_{\frac{1}{\lambda}\varphi}\big)(Bu^m + v^m). \end{cases} \tag{4.4}$$

Similarly, we derive the following iterative algorithm for TVL1L2 model:

$$\begin{cases} u^m = A^{-1}(g-\lambda B^T v_1^m - \lambda W^T v_2^m), \\ {v_1}^{m+1} = \kappa v_1^m + (1-\kappa)\big(I - prox_{\frac{1}{\lambda}\varphi_1}\big)(Bu^m + v_1^m), \\ {v_2}^{m+1} = \kappa v_2^m + (1-\kappa)\big(I - prox_{\frac{1}{\lambda}\varphi_2}\big)(Wu^m + v_2^m). \end{cases} \tag{4.5}$$

At the end of this part, we follow notations of [30] and give a fast algorithm for computing u^m in (4.4). Let $v^m = (w_1^m; w_2^m)$ with $w_i^m \in \mathbb{R}^{n^2}$, $i = 1,2$, then $u^m = A^{-1}(g - \lambda B^T v^m)$ can be written as

$$(K^T K + \alpha I) u^m = K^T f - \lambda {D^{(1)}}^T w_1^m - \lambda {D^{(2)}}^T w_2^m, \tag{4.6}$$

where α is a regularization parameter discussed before. Under the periodic boundary condition, $K^T K$ is block circulation. Therefore, the matrix on the left-hand side of (4.6) can be diagonalized by two-dimensional discrete Fourier transform $\mathcal{F}$. Using the convolution theorem of Fourier transform, we get

$$u^m = \mathcal{F}^{-1}\Big(\frac{\mathcal{F}(K)^* \mathcal{F}(f) - \lambda \mathcal{F}(D^{(1)})^* \circ \mathcal{F}(w_1^m) - \lambda \mathcal{F}(D^{(2)})^* \circ \mathcal{F}(w_2^m)}{\alpha I + \mathcal{F}(K)^* \circ \mathcal{F}(K)}\Big) \triangleq \mathcal{F}^{-1}(w^m). \tag{4.7}$$

where "$*$" denotes complex conjugate, "$\circ$" denotes component-wise multiplication, and the division is component-wise as well. For more details please refer to [30]. The method can also be applied to scheme (4.5) for fast algorithm.

4.2 The relationship with other algorithms

Bermúdez and Moreno derived their results for solving variational inequalities by the Yosida approximation of the subdifferential of φ in 1981. Recently, J. F. Aujol [17] used their results to deal with the TV related problems, such as the TV denoising, smoothed TV-based image restoration, and get a class of iterative algorithms. He also discovered that the Chambolle's projection algorithm can be viewed as a particular instance of his proposed algorithm. In Aujol's paper, numerous numerical examples show that such iterative algorithms are comparable with a general class of algorithm introduced by Nestorv. Recently, C. A. Micchelli proposed a proximity point algorithm for TV denoising named FP^{2}O which also includes the Chambolle's projection algorithm as a special case of it. We have found that both of them can be viewed as generation of the Chambolle's projection algorithm but from different points of view. The FP^{2}O algorithm is based on the Picard iteration, and the Bermúdez-Moreno algorithm is to extend the operator A from identity operator to symmetric positive operator. Since we consider these two points, in some degree our algorithms are a more generation for both of the FP^{2}O and the Bermúdez-Moreno algorithms.

Now, let us firstly show that the Bermúdez-Moreno algorithm is the particular case of scheme (4.4) when $\kappa=0$. Denote

$$H=\partial_\varphi, \quad L_\mu=\Big(I+\frac{1}{\mu}H\Big)^{-1} \quad \text{and} \quad H_\mu=\mu(I-L_\mu),$$

then the Bermúdez-Moreno iterative scheme is

$$\left\{\begin{array}{l} x^m=A^{-1}(g-\lambda B^T y^m), \\ y^{m+1}=H_\mu\Big(Bx^m+\dfrac{1}{\mu}y^m\Big). \end{array}\right. \tag{4.8}$$

Let $y^m/\mu=v^m$, $x^m=u^m$, then the Bermúdez-Moreno iterative scheme can be written equivalently as

$$\left\{\begin{array}{l} u^m=A^{-1}(g-\lambda B^T v^m), \\ v^{m+1}=(I-L_\mu)(Bu^m+v^m). \end{array}\right. \tag{4.9}$$

The remainder is to show that $L_\mu=prox_{\varphi/\lambda}$. Indeed, for any convex function φ, its proximity point u at any fixed point v is equal to the minimizer of $\|x-v\|^2/2+\varphi(x)/\lambda$. Since such objective function is strictly convex, we conclude that

$$\partial\Big\{\frac{1}{2}\|x-v\|^2+\frac{1}{\lambda}\varphi(x)\Big\}\Big|_{x=u}=0,$$

i.e.,

$$u-v+\frac{1}{\lambda}\partial_\varphi(u)=0.$$

It follows that

$$u=\Big(I+\frac{1}{\lambda}H\Big)^{-1}(v)=L_\lambda(v),$$

which validates the equality of the operators L_μ and $prox_{\varphi/\lambda}$. Furthermore,

- If we choose $\kappa=0$ and $A=I$ in the scheme (4.4), or choose $A=I$ in the Bermúdez-Moreno scheme, then we get the Jia-Zhao denoising algorithm.
- If we only choose $A=I$ in the scheme (4.4), then the scheme (4.4) reduces to the FP^{2}O algorithm.
- If we only choose $\kappa=0$ in the scheme (4.4), then the scheme (4.4) reduces to the Bermúdez-Moreno algorithm.

Since the connection among the split Bregman algorithm, the Jia-Zhao denoising algorithm, and the FP^{2}O algorithm has been discussed in several papers [11, 18, 31], we do not repeat it here. But we have determined in the introduction to compare our algorithms with the split Bregman algorithm. To this end, we recall that the split Bregman algorithm applied to the TVL2 and TVL1L2 models which have the following forms respectively:

$$\begin{cases} u^{m+1} = \mathrm{argmin}_u \|Ku-f\|_2^2 + \lambda\|Bu-d^m+b^m\|_2^2, \\ d^{m+1} = \mathrm{argmin}_d \varphi(d) + \dfrac{\lambda}{2}\|Bu^{m+1}-d+b^m\|_2^2, \\ b^{m+1} = b^m + (Bu^{m+1}-d^{m+1}), \end{cases} \tag{4.10}$$

and

$$\begin{cases} u^{m+1} = \mathrm{argmin}_u \|Ku-f\|_2^2 + \lambda\|Bu-d^m+b_1^m\|_2^2 + \lambda\|Bu-w^m+b_2^m\|_2^2, \\ d^{m+1} = \mathrm{argmin}_d \varphi_1(d) + \dfrac{\lambda}{2}\|Bu^{m+1}-d+b_1^m\|_2^2, \\ w^{m+1} = \mathrm{argmin}_w \varphi_2(w) + \dfrac{\lambda}{2}\|Wu^{m+1}-w+b_2^m\|_2^2, \\ b_1^{m+1} = b_1^m + (Bu^{m+1}-d^{m+1}), \\ b_2^{m+1} = b_2^m + (Wu^{m+1}-w^{m+1}). \end{cases} \tag{4.11}$$

5 Convergence analysis

The following lemma plays a central role in our convergence analysis.

Lemma 5.1 (Opial κ-averaged property [19]). *If C is a closed and convex set in $\mathbb{R}^{n^2}$ and $P: C \to C$ has at least one fixed point and is a nonexpansive mapping, i.e., for any x,y it holds that $\|P(x)-P(y)\|_2 \leq \|x-y\|_2$; then for $\kappa \in (0,1)$, P_κ is nonexpansive, maps C to itself, and has the same set of the fixed points as P. Furthermore, for any $u \in C$ and $\kappa \in (0,1)$, the Picard sequence of P_κ converges to a fixed point of P.*

Under some assumptions, we shall show that both of S and $\bar{S}$ fulfill the conditions required by Lemma 5.1 which can guarantee the schemes (4.4) and (4.5) converge to the fixed point of S and $\bar{S}$ respectively. The existence of the fixed point of our discussed operators is a direct result of the existence of solution to the strictly convex problems. If we can construct a closed and convex set containing at least one fixed point of S or $\bar{S}$, and at the same time this set is the range of the operators, then we validate the first condition. Therefore, we only need to discuss under what assumptions the nonexpansive condition can be satisfied. As a guidance, we list the main points we will talk about.

- Determining an upper bound for the parameter λ and showing under such an upper bound the operators S and $\bar{S}$ must be nonexpansive.
- Constructing a closed and convex set in which at least one fixed point of the operators lies, and at the same time this set is the range of the operators.

Proposition 5.1. If φ is a convex function, and λ is a positive number such that $\|I-\lambda BA^{-1}B^T\|_2 \leq 1$, where $\|\cdot\|_2$ is the spectral norm and represents the biggest eigenvalue of some matrix, then S is nonexpansive.

The proof of proposition above is similar to Lemma 3.3 in [11]. We do not repeat their arguments here. From practical point of view, the parameter λ is not easy to be determined by the above conclusion. The following corollary would be a good complementarity for such a drawback since it only depends on the regularization and the spectral norm of BB^T whose eigenvalues have been determined in [11].

Corollary 5.1. *If φ is a convex function, λ is a positive number such that $\lambda \leq 2\alpha/\|BB^T\|_2$, then S is nonexpansive.*

Proof. Since $D^{(1)}$, $D^{(2)}$, $K^TK+\alpha I$ can be diagonalized by two-dimensional discrete Fourier transform, it follows that $D^{(i)}A^{-1}=A^{-1}D^{(i)}$, $i=1,2$. Thus

$$BA^{-1}=(D^{(1)};D^{(2)})A^{-1}=diag(A^{-1},A^{-1})B.$$

Therefore,

$$\begin{aligned}\|I-\lambda BA^{-1}B^T\|_2 &= \|I-\lambda diag(A^{-1},A^{-1})BB^T\|_2 \\ &= \begin{cases} \lambda\lambda_N-1, & \text{if } 1-\lambda\lambda_1<\lambda\lambda_N-1, \\ 1-\lambda\lambda_1, & \text{otherwise,}\end{cases}\end{aligned} \tag{5.1}$$

where λ_1, λ_N are the smallest and the biggest eigenvalues of $diag(A^{-1},A^{-1})BB^T$ respectively. In the first case,

$$\begin{aligned}\lambda\lambda_N-1 &= \lambda\|diag(A^{-1},A^{-1})BB^T\|_2-1 \\ &\leq \lambda\|diag(A^{-1},A^{-1})\|_2\|BB^T\|_2-1 \\ &\leq \lambda\alpha^{-1}\|BB^T\|_2-1.\end{aligned} \tag{5.2}$$

The last inequality is valid because of that $\|diag(A^{-1},A^{-1})\|_2=\|A^{-1}\|_2$ and $\|A^{-1}\|_2=\|(K^TK+\alpha I)^{-1}\|\leq\alpha^{-1}$. By the condition $\lambda\leq 2\alpha/\|BB^T\|_2$, we conclude that $\lambda\lambda_N-1\leq 1$ which implies $\|I-\lambda BA^{-1}B^T\|_2\leq 1$, so S is nonexpansive by Proposition 5.1. In the second case, $\|I-\lambda BA^{-1}B^T\|_2=1-\lambda\lambda_1\leq 1$ always holds since $\lambda\lambda_1$ is a positive number. Therefore, $\lambda\leq 2\alpha/\|BB^T\|_2$ is sufficient to guarantee S to be nonexpansive. $\square$

Before deducing similar results for operator $\bar{S}$, we denote

$$Y=\begin{pmatrix} I-\lambda BA^{-1}B^T & -\lambda BA^{-1}W^T \\ -\lambda WA^{-1}B^T & I-\lambda WA^{-1}W^T\end{pmatrix}. \tag{5.3}$$

Lemma 5.2 (see [11]). *If φ is a convex function, then operator $I-prox_{\varphi/\lambda}$ is nonexpansive.*

Proposition 5.2. If φ_i, $i=1,2$ are convex functions, and λ is a positive number such that $\|Y\|_2\leq 1$, then the operator $\bar{S}$ is nonexpansive.

Proof. Since for any $u_i\in\mathbb{R}^{2n^2}$, $i=1,2$ and $v_i\in\mathbb{R}^{n^2}$, $i=1,2$, we derive that

$$\begin{aligned}
&\left\|\bar{S}\circ\begin{pmatrix}u_1\\u_2\end{pmatrix}-\bar{S}\circ\begin{pmatrix}v_1\\v_2\end{pmatrix}\right\|_2^2\\
=&\left\|\begin{pmatrix}(I-prox_{\frac{1}{\lambda}\varphi_1})\circ L_1\circ(u_1;u_2)-(I-prox_{\frac{1}{\lambda}\varphi_1})\circ L_1\circ(v_1;v_2)\\(I-prox_{\frac{1}{\lambda}\varphi_2})\circ L_2\circ(u_1;u_2)-(I-prox_{\frac{1}{\lambda}\varphi_2})\circ L_2\circ(v_1;v_2)\end{pmatrix}\right\|_2^2\\
=&\sum_{i=1,2}\left\|(I-prox_{\frac{1}{\lambda}\varphi_i})\circ L_i\circ(u_1;u_2)-(I-prox_{\frac{1}{\lambda}\varphi_i})\circ L_i\circ(v_1;v_2)\right\|_2^2\\
\leq&\sum_{i=1,2}\|L_i\circ(u_1;u_2)-L_i\circ(v_1;v_2)\|_2^2=\left\|\begin{pmatrix}L_1\circ(u_1;u_2)-L_1\circ(v_1;v_2)\\L_2\circ(u_1;u_2)-L_2\circ(v_1;v_2)\end{pmatrix}\right\|_2^2\\
=&\left\|Y\begin{pmatrix}u_1\\u_2\end{pmatrix}-Y\begin{pmatrix}v_1\\v_2\end{pmatrix}\right\|_2^2\leq\|Y\|_2^2\left\|\begin{pmatrix}u_1\\u_2\end{pmatrix}-\begin{pmatrix}v_1\\v_2\end{pmatrix}\right\|_2^2,
\end{aligned}\tag{5.4}$$

where the second and the third equations are based on the Pythagorean identity and the first inequality on Lemma 5.2. Therefore, we conclude that if λ is a positive number such that $\|Y\|_2\leq 1$, then the operator $\bar{S}$ is nonexpansive. ☐

Note that

$$\begin{aligned}
Y=&I-\lambda\begin{pmatrix}BA^{-1}&BA^{-1}\\WA^{-1}&WA^{-1}\end{pmatrix}\begin{pmatrix}B^T&0\\0&W^T\end{pmatrix}\\
=&I-\lambda\begin{pmatrix}diag(A^{-1},A^{-1})B&diag(A^{-1},A^{-1})B\\A^{-1}W&A^{-1}W\end{pmatrix}\begin{pmatrix}B^T&0\\0&W^T\end{pmatrix}\\
=&I-\lambda diag(A^{-1},A^{-1},A^{-1})\begin{pmatrix}B&B\\W&W\end{pmatrix}\begin{pmatrix}B^T&0\\0&W^T\end{pmatrix}\\
=&I-\lambda diag(A^{-1},A^{-1},A^{-1})\hat{B}\hat{B}^T,
\end{aligned}\tag{5.5}$$

where $\hat{B}=(B;W)$. We have to mention that in (5.5) we have assumed $WA^{-1}=A^{-1}W$. This equality is not always valid for general wavelet frame except in some special cases such as W is a convolution operator. Repeating the arguments in the proof of Corollary 5.1, we get the following corollary for operator $\bar{S}$.

Corollary 5.2. *If φ is a convex function and the frame W is a convolution operator, λ is a positive number such that $\lambda\leq 2\alpha/\|\hat{B}\hat{B}^T\|_2$ with $\hat{B}=(B;W)$, then $\bar{S}$ is nonexpansive.*

We construct a closed and convex set for Lipschitz continuous convex function φ; that is

$$C_\varphi:=\left\{z:z\in\mathbb{R}^{2n^2},\ \|z\|_2\leq Lip_\varphi/\lambda\right\},\tag{5.6}$$

where Lip_φ is the Lipschitz constant of φ, given by the smallest constant L such that $|\varphi(u)-\varphi(v)| \le L\|u-v\|_2$. For Lipschitz continuous convex functions φ_1, φ_2, we define $C_{1,2} := C_{\varphi_1} \times C_{\varphi_2}$, where "$\times$" is the Descartes production. By the Tychonov theorem, $C_{1,2}$ is a compact and hence closed set since it lies in a finite dimensional space. The convexity is obvious by the definition of convex set.

Proposition 5.3. If φ is a Lipschitz continuous convex function, then S maps $\mathbb{R}^{2n^2}$ into C_φ.

Proposition 5.4. If φ_1, φ_2 are Lipschitz continuous convex functions, then $\bar{S}$ maps $\mathbb{R}^{3n^2}$ into $C_{1,2}$.

The validity of propositions 5.3 and 5.4 depends on an inclusion: $(I-prox_{\varphi/\lambda})v \in C_\varphi$ for any variable v in the defined domain, since both S and $\bar{S}$ are compositions of an affine transform and the operator $(I-prox_{\varphi/\lambda})$. The inclusion has been shown in [11], so we do not prove Propositions 5.3 and 5.4 at length. But here we would like to emphasize a fact that C_φ and $C_{1,2}$ are exactly the sets we want since they are the range of operator S and $\bar{S}$ respectively and each of them contains at least one fixed point. At last, we have the following results as direct consequences of Lemma 3.1:

Proposition 5.5. If φ is a Lipschitz continuous convex function, $0<\kappa<1$, and λ is a positive number such that $\|I-\lambda BA^{-1}B^T\|_2 \le 1$ or $\lambda \le 2\alpha/\|BB^T\|_2$, then $\{u^m\}$ and $\{v^m\}$ in the scheme (4.4) converge to the solution of (2.7) and the fixed point of S.

Proposition 5.6. If φ_1, φ_2 are Lipschitz continuous convex functions, $0<\kappa<1$, and λ is a positive number such that $\|Y\|_2 \le 1$, then $\{u^m\}$ and $\{v^m\}$ in the scheme (4.5) converge to the solution of (2.8) and the fixed point of $\bar{S}$.

J. F. Aujol in [17] has shown the convergence under the assumption $\lambda \le 2\alpha/\|BB^T\|_2$ for the Bermúdez-Moreno algorithm; this is just the case $\kappa=0$ of the scheme (4.4). Together with their result, we get a more complete convergent result for the scheme (4.4).

Proposition 5.7. If φ is a Lipschitz continuous convex function, $0 \le \kappa < 1$, and λ is a positive number such that $\lambda \le 2\alpha/\|BB^T\|_2$, then $\{u^m\}$ and $\{v^m\}$ in the scheme (4.4) converge to the solution of (2.7) and the fixed point of S.

Notice that the conditions in Proposition 5.5 are different from those in Proposition 5.7. In the former proposition, the interval of κ excludes the case $\kappa=0$ corresponding to the Bermúdez-Moreno scheme; in the later one, though it includes such a special case, the upper bound of λ has been tighten.

6 Numerical experiments

In this section, we report some numerical experiments with the proposed first-order algorithms, and compare the computational performance of our algorithms with that of the

Figure 1: The classical Lena and Cameraman images with sizes 256×256.

split Bregman algorithm, which is one of the best algorithms for solving TV-based variation models at present. All the experiments were performed using MATLAB (version 7.01), on a computer equipped with an Intel Pentium-IV 3.0GHz processor, with 512MB of RAM, and running Windows XP.

We used two images of Cameraman and Lena with sizes 256×256 as the original images for our experiments to illustrate our study. These images are shown in Fig. 1. The observed images are modeled as

$$f = Ku + \eta, \eta \sim N(0,\sigma^2), \tag{6.1}$$

with $N(0,\sigma^2)$ being Gaussian noise, and K represents a blurring or convolution operator. We consider five benchmark deblurring problems [32] to test the scheme (4.4), summarized in Table 1. We only use experiment 1 to test the scheme (4.5) because our aim is to show the proposed first-order algorithm can deal with the complicated TVL1L2 model and performs better than the split Bregman method. We choose the Haar wavelet with level 2 as wavelet analysis. The quality of restoration images $\tilde{u}$ obtained from algorithms is evaluated by the signal-to-noise ratio

$$SNR := 10\log\left\{\frac{\|f-\bar{u}\|}{\|f-\tilde{u}\|}\right\}, \tag{6.2}$$

Table 1: Details of the images restoration experiments.

Experiments	Operator K	σ^2
1	9×9 uniform	0.56^2
2A	Gaussian	2
2B	Gaussian	8
3A	$h_{ij}=1/(1+i^2+j^2)$	2
3B	$h_{ij}=1/(1+i^2+j^2)$	8

Table 2: Numerical results for Lena image, the pair $(\cdot,\cdot)$ is used to report both the SNR value (the first number) and the number of CPU time (seconds) (the second number).

Experiments	$1(\beta=0.08)$	$2A(\beta=0.50)$	$2B(\beta=2.00)$	$3A(\beta=0.10)$	$3B(\beta=0.60)$
Scheme (4.10)	(13.76,8.98)	(25.36,15.66)	(15.56,18.05)	(15.60,10.17)	(11.74,14.33)
Scheme (4.4)	(13.90,6.63)	(25.44,12.30)	(15.80,12.33)	(15.65,7.59)	(11.78,10.59)

Table 3: Numerical results for Cameraman image, the pair $(\cdot,\cdot)$ is used to report both the SNR value (the first number) and the number of CPU time (seconds) (the second number).

Experiments	$1(\beta=0.06)$	$2A(\beta=0.40)$	$2B(\beta=2.00)$	$3A(\beta=0.08)$	$3B(\beta=0.80)$
Scheme (4.10)	(15.16,13.10)	(27.41,15.80)	(17.71,17.50)	(15.39,10.45)	(12.22,14.48)
Scheme (4.4)	(15.44,8.32)	(27.46,11.30)	(17.74,12.25)	(15.69,7.01)	(12.24,11.02)

where $\bar{u}$ represents the mean of restored image. Since isotropic total variation model always perform better than anisotropic total variation model, we adopt isotropic total variation in experiments, but the results are similar for anisotropic total variation model. Therefore, we choose $\varphi(\cdot)=\beta\psi(\cdot)$ with β being a regularization parameter in schemes (4.4) and (4.10), and $\varphi_1(\cdot)=\beta_1\psi_1(\cdot)$, $\varphi_2(\cdot)=\beta_2\psi_2(\cdot)$ with β_1, β_2 being regularization parameters in schemes (4.5) and (4.11). We tune these parameters in each case for best improvement in SNR, so that the comparison is carried out in the regime that is relevant in practice. We fixed $\kappa=0.0001$, $\alpha=0.002$ and $\lambda=0.0005$ (i.e., $\lambda=\alpha/4$) by Corollary 5.1 and discussion in the convergence analysis.

Since our proposed algorithms are computing a fixed point of Picard iteration sequence, we stop the algorithms when the iteration sequence has little change. Therefore, for the scheme (4.4) iterations are terminated when the following condition is satisfied

$$\frac{\|v^{m+1}-v^m\|_2}{\|v^{m+1}\|_2}\leq TOL, \tag{6.3}$$

where TOL denotes a prescribed tolerance value. For the split Bregman algorithm, i.e., scheme (4.10), we terminated the iterations when the iteration sequence $\{b^m\}$ satisfies

$$\frac{\|b^{m+1}-b^m\|_2}{\|b^{m+1}\|_2}\leq TOL. \tag{6.4}$$

In our experiments, we set TOL=0.005. For schemes (4.5) and (4.11), similar stopping criterions were used. Our results for comparisons between schemes (4.4) and (4.10) are reported in Table 2, Table 3 and Fig. 2. From there, one can see that the new scheme cost less time but get higher SNR than that of scheme (4.10). Fig. 3 shows the results of schemes (4.5) and (4.11) that illustrates the new schemes perform better than the split Bregman method.

Figure 2: Image restoration results for comparisons between the scheme (4.4) and the split Bregman method. (top): Degraded images (convolved with a $h_{ij}=1/(1+i^2+j^2)$ kernel, and then degraded by a zero mean Gaussian noise with derivation $\sigma=8$); (middle): Restored images by the scheme (4.4); (bottom): Restored images by the scheme (4.10).

7 Conclusions

In this paper, we propose new first-order schemes for total variation (TV) model and also for wavelet analysis with TV model. Under the proximity point operator framework, these schemes are natural generations of the proximity point algorithm for total variation denoising. We firstly derive fixed point formulations for the models, and then design new first-order algorithms by introducing the Picard sequence. We proved the convergence of the new algorithms based on the Opial κ-averaged theorem. Moreover, a series of

Figure 3: Image restoration results for comparisons between the scheme (4.5) and the split Bregman method with $\beta_1=\beta_2=0.04$. (top): Degraded images (convolved with a 9×9 uniform kernel, and then degraded by a zero mean Gaussian noise with derivation $\sigma=0.56^2$); (middle): Restored images by the scheme (4.5); (middle left): SNR=15.39dB, time=31.37s; (middle right): SNR=15.83dB, time=34.61s; (bottom): Restored images by the scheme (4.11); (bottom left): SNR=15.16dB, time=33.64s; (bottom right): SNR=15.56dB, time=39.08s.

connection among current algorithms, such as Chambolle's projection algorithm, the split Bregman algorithm, the Bermúdez-Morenoalgorithm, the Jia-Zhao denoising algorithm, and the fixed point algorithm, have been discovered; all of them are famous because either of their history or efficiency. At last, we make comparisons with the split Bregman algorithm which is one of the best algorithms for solving TV-based variation models at present. Numerical experiments illustrate that the proposed algorithms perform better than the split Bregman algorithm.

Acknowledgments

This research is supported by the National Science Foundation of China under grants (No. 61271014), Specialized Research Fund for the Doctoral Program of Higher Education (20124301110003) and the Graduated Students Innovation Fund of Hunan Province (CX2012B238).

References

[1] L. RUDIN, S. OSHER, AND E. FATEMI, *Nonlinear total variation based noise removal algorithms*, Phys. D, 60 (1992), pp. 259–268.

[2] G. AUBERT AND P. KORNPROBST, *Mathematical problems in image processing*, Volume 147 of Applied Mathematical Sciences, Springer-Verlag, 2002.

[3] T. CHAN AND J. SHEN, *Image processing and analysis-Variational*, PDE, Wavelet, and Stochastic Methods, SIAM Publisher, 2005.

[4] A. CHAMBOLLE, *An algorithm for total variation minimization and applications*, J. Math. Imag. Vision, 20 (2004), pp. 89–97.

[5] A. CHAMBOLLE, *Total variation minimization and a class of binary MRF models*, In EMMCVPR 05, Volume 3757 of Lecture Notes in Computer Sciences, pp. 136–152, 2005.

[6] T. GOLDSTEIN AND S. OSHER, *The split Bregman method for regularization problems*, SIAM J. Imag. Sci., 2 (2009), pp. 323–343.

[7] X. ZHANG, M. BURGER, AND S. OSHER, *A unified primal-dual algorithm framework based on Bregman iteration*, J. Sci. Comput., 46(1) (2011), pp. 20–46.

[8] Y. NESTEROV, *Smooth minimization of non-smooth functions,* Math. Program. A, 103(1) (2005), pp. 127–152.

[9] P. WEISS, G. AUBERT, AND L. BLANC-FERAUD, *Efficient schemes for total variation minimization under constraints in image processing,* SIAM J. Sci. Comput., 31(3) (2010), pp. 2047–2080.

[10] A. BERMÚDEZ AND C. MORENO, *Duality methods for solving variational inequalities,* Comput. Math. Appl., 7 (1981), pp. 43–58.

[11] C. A MICCHELLI, L. SHEN, AND Y. XU, *Proximity algorithms for image models: denoising,* Inverse Problem, 27 (2011), pp. 045009.

[12] J. F. CAI, S. OSHER AND Z. SHEN, *Split Bregman methods and frame based image restoration,* Multiscale Model. Simul., 8(2) (2009), pp. 337–369.

[13] T. GOLDSTEIN, X. BRESSON, AND S. OSHER, *Geometric applications of the split Bregman method: segmentation and surface reconstruction*, J. Sci. Comput., 45(1-3) (2010), pp. 272–293.

[14] M. GU, L. H. LIM, AND C. J. WU, *PARNES: a rapidly convergent algorithm for accurate recovery of sparse and approximately sparse signal,* arXiv: 0911.0492, 2009.

[15] S. BECKER, J. BOBIN, AND E. J. CANDÈS, *Nesta: a fast and accurate first-order method for sparse recovery,* arXiv: 0904.3367, 2009.

[16] A. BECK AND M. TEBOULLE, *A fast iterative shrinkage-thresholding algorithm with application to wavelet-based image deblurring acoustics, speech, and signal processing,* IEEE Int. Conference on, 10 (2009), pp. 693–696.

[17] J. F. AUJOL, *Some first-order algorithms for total variation based image restoration,* J. Math. Imag. Vision, 34(3) (2010), pp. 307–327.

[18] R. Q. JIA AND H. ZHAO, *A fast algorithm for the total variation model of image denoising,* Adv. Comput. Math., 33 (2010), pp. 231–241.

[19] Z. OPIAL, *Weak convergence of the sequence of successive approximations for nonexpansive mapping,* Bulletin American Mathematical Society, 73 (1967), pp. 591–597.
[20] M. V. AFONSO, J. M. BIOUCAS-DIAS, AND M. FIGUEIREDO, *Fast image recovery using variable splitting and constrained optimization,* IEEE Trans. Image Process., 19(9) (2010), pp. 2345–2356.
[21] L. HE, T. C. CHANG AND S. OSHER ET AL, *MR image reconstruction by using the iterative refinement method and nonlinear inverse scale space methods,* UCLA CAM Report(06-35), 2006.
[22] J. YANG, Y. ZHANG, AND W. YIN, *A fast TVL1-L2 minimization algorithm for signal reconstruction from partial Fourier data,* IEEE Journal of Selected Topics in Signal Processing, 4 (2009), pp. 288–297.
[23] H. ZHANG, L. Z. CHENG, AND J. P. LI, *Reweighted minimization model for MR image reconstruction with split bregman method,* Sci. China Ser., 54(1) (2011), pp. 1–10.
[24] S. BECKER, E. J. CANDÈS, AND M. GRANT, *Templates for convex cone problems with applications to sparse signal recovery,* arXiv: 1009.2065, 2010.
[25] J. MOREAU, *Proximitéet dualitédans un espace hilbertien,* Bull. Soc. Math. France, 93 (1965), pp. 273–299.
[26] P. COMBETTES AND V. WAJS, *Signal recovery by proximal forward-backward splitting,* Multiscale Model. Simul., 4 (2005), pp. 1168–1200.
[27] P. COMBETTES AND J. PESQUET, *A proximity decomposition method for solving convex variational inverse problem,* Inverse Problem, 24(6) (2008), pp. 65014–65040.
[28] P. COMBETTES AND J. PESQUET, *A Douglas-Rachford splitting approach to nonsmooth convex variational signal recovery,* IEEE Journal of Selected Topics in Signal Processing, 1 (2007), pp. 564–574.
[29] D. BERTSEKAS, A. NEDIĆ, AND A. E. OZDAGLAR, Convex Analysis and Optimization, Athena Scientific and Tsinghua University Press, pp: 227, 2006.
[30] Y. WANG, J. YANG, W. YIN, AND Y. ZHANG, *A new alternating minimization algorithm for total variation image reconstruction,* SIAM J. Imag. Sci., 1 (2008), pp. 248–272.
[31] D. CHENG, H. ZHANG, AND L. CHENG, *A fast fixed point algorithm for total variation deblurring and segmentation,* J. Math. Imag. Vision, 43(3) (2012), pp. 167–179.
[32] M. FIGUEIREDO AND R. NOWAK, *An EM algorithm for wavelet-based image restoration,* IEEE Trans. Image Process., 12 (2003), pp. 906–916.